Data Collection and Preprocessing Phase

Date	17 April 2024
Team ID	738184
Project Title	Eye Disease Detection using Deep Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan

Section	Description
Project Overview	This project focuses on developing a deep learning model for eye disease detection. The goal is to accurately classify images into categories like normal, cataract, glaucoma, and diabetic retinopathy.
Data Collection Plan	Data for this project is collected from Kaggle which provides various medical imaging datasets related to eye diseases.
Raw Data Sources Identified	The raw data source for this project include datasets obtained from Kaggle, the popular platform for data science competitions and datasets. The provided dataset contains a set of images with eye diseases in respective sub folders: Cataract, Diabetic Retinopathy, Glaucoma and normal each of different pixel sizes.

Raw Data Sources

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset	The dataset consists of Normal, Diabetic Retinopathy, Cataract and Glaucoma retinal images where each class has approximately 1000 images. These images are collected from various sources like IDRiD, Oculus recognition, HRF etc..	https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification/data	IMG (.png, .jpg, .jpeg)	772.22 MB	Public